■ SPINE

Neurological recovery after early versus delayed surgical decompression for acute traumatic spinal cord injury

A PROSPECTIVE, OBSERVATIONAL, EUROPEAN MULTICENTRE (SCI-POEM) COHORT STUDY

A. J. F. Hosman, G. Barbagallo, the SCI-POEM Study Group, J. J. van Middendorp

From Radboud University Medical Center, Nijmegen, Netherlands

Aims

The aim of this study was to determine whether early surgical treatment results in better neurological recovery 12 months after injury than late surgical treatment in patients with acute traumatic spinal cord injury (tSCI).

Methods

Patients with tSCI requiring surgical spinal decompression presenting to 17 centres in Europe were recruited. Depending on the timing of decompression, patients were divided into early (≤ 12 hours after injury) and late (> 12 hours and < 14 days after injury) groups. The American Spinal Injury Association neurological (ASIA) examination was performed at baseline (after injury but before decompression) and at 12 months. The primary endpoint was the change in Lower Extremity Motor Score (LEMS) from baseline to 12 months.

Results

The final analyses comprised 159 patients in the early and 135 in the late group. Patients in the early group had significantly more severe neurological impairment before surgical treatment. For unadjusted complete-case analysis, mean change in LEMS was 15.6 (95% confidence interval (CI) 12.1 to 19.0) in the early and 11.3 (95% CI 8.3 to 14.3) in the late group, with a mean between-group difference of 4.3 (95% CI -0.3 to 8.8). Using multiply imputed data adjusting for baseline LEMS, baseline ASIA Impairment Scale (AIS), and propensity score, the mean between-group difference in the change in LEMS decreased to 2.2 (95% CI -1.5 to 5.9).

Conclusion

Compared to late surgical decompression, early surgical decompression following acute tSCI did not result in statistically significant or clinically meaningful neurological improvements 12 months after injury. These results, however, do not impact the well-established need for acute, non-surgical tSCI management. This is the first study to highlight that a combination of baseline imbalances, ceiling effects, and loss to follow-up rates may yield an overestimate of the effect of early surgical decompression in unadjusted analyses, which underpins the importance of adjusted statistical analyses in acute tSCI research.

Cite this article: *Bone Joint J* 2023;105-B(4):400–411.

Introduction

Traumatic spinal cord injury (tSCI) has a profound impact on patients' physical and psychosocial wellbeing. Despite the low frequency, at 15 to 39 per million population worldwide, this devastating injury imposes a substantial burden on healthcare systems.[1] While many different treatments have been investigated, none have been shown in high-quality studies to convincingly improve neurological or functional recovery.[2] Over a century ago the American surgeon, Dr Herbert Leslie Burrell (1856 to 1910) postulated that the therapeutic role of acute surgical decompression, based on the assumption "that if pressure on the cord is allowed to remain for many hours, irreparable damage to the cord may take place."[3]

Correspondence should be sent to A. J. F. Hosman; email: allard.hosman@radboudumc.nl

doi:10.1302/0301-620X.105B4.BJJ-2022-0947.R2 $2.00

Bone Joint J 2023;105-B(4):400–411.

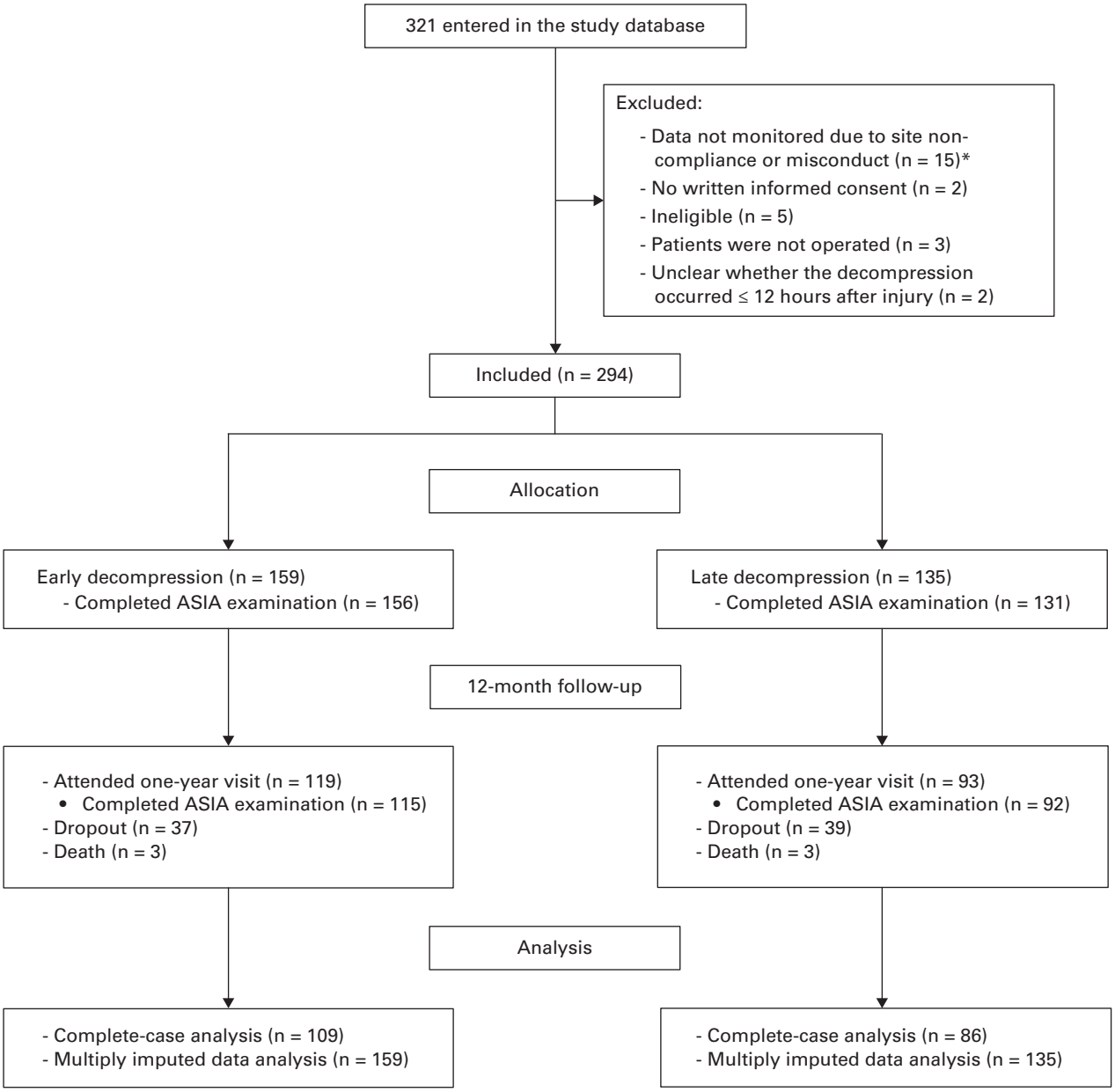

Fig. 1

Study flowchart. *This included all 13 patients recruited from two study sites. As such, the final enrolled patients were recruited from 15 out of the original 17 study centres. ASIA, American Spinal Injury Association.

However, despite the positive effects of acute spinal decompressions reported in preclinical studies,[4] the strong biological rationale for early intervention has not translated into neurological benefits that could be consistently seen in any of the human tSCI studies published so far.[5,6]

The challenge presented by the paucity of robust, high-quality studies was reflected in the most recent international clinical practice guidelines, which recommended that early surgery be offered as a "treatment option" but acknowledged that the recommendation was "weak" and the quality of evidence was "low".[5] No further well-powered, comparative studies have been produced since the publication of these guidelines. Nonetheless, the "time is spine" adage, which implies that early surgical decompression of the spinal cord results in improved outcomes following tSCI, has been widely adopted among surgeons involved in acute tSCI management over the last decade.[6-8] From a public health policy-making perspective, it is important to consider the various medicolegal and operational feasibility challenges associated with treating acute tSCI, as well as the non-trivial infrastructural and systemic changes required for implementing early surgical management practices.[9] These practical realities, and the fact that the guidelines

Table I. Summary of patient characteristics at baseline by treatment group. Results are presented using patients with data available for individual variables; no imputation was performed to impute missing data.

Characteristic	Total	Early decompression	Late decompression	p-value
Patients, n	294	159	135	
Mean age, yrs (SD)	46.7 (17.9)	44.7 (17.1)	49.1 (18.6)	0.049*
Sex, n (%)	294	159	135	0.998†
Male	233 (79.3)	126 (79.2)	107 (79.3)	
Female	61 (20.7)	33 (20.8)	28 (20.7)	
Prior admission to another hospital after injury, n (%)	294	159	135	< 0.001†
Yes	131 (44.6)	37 (23.3)	94 (69.6)	
No	163 (55.4)	122 (76.7)	41 (30.4)	
CCI, n (%)	292	159	133	0.005‡
0	254 (87.0)	147 (92.5)	107 (80.5)	
1	23 (7.9)	6 (3.8)	17 (12.8)	
2	12 (4.1)	5 (3.1)	7 (5.3)	
3	2 (0.7)	0	2 (1.5)	
6	1 (0.3)	1 (0.6)	0	
Injury Severity Score, n	282	155	127	0.002*
Mean (SD)	17.6 (9.5)	19.1 (9.9)	15.6 (8.7)	
ASIA neurological examinations	284	153	131	< 0.001*
Mean time between injury and ASIA examination at baseline, hrs (SD)	11.5 (26.9)	3.2 (2.3)	21.3 (37.4)	
AIS grade, n (%)	283	153	130	< 0.001†
A	110 (38.9)	72 (47.1)	38 (29.2)	
B	48 (17.0)	35 (22.9)	13 (10.0)	
C	41 (14.5)	19 (12.4)	22 (16.9)	
D	84 (29.7)	27 (17.6)	57 (43.8)	
Total motor score, n	268	147	121	0.015*
Mean (SD)	51.2 (28.7)	47.3 (27.0)	55.9 (30.1)	
LEMS, n	275	151	124	< 0.001*
Mean (SD)	15.5 (19.5)	10.1 (16.6)	22.0 (20.8)	
Light touch score, n	275	151	124	< 0.001*
Mean (SD)	74.3 (31.5)	69.1 (30.6)	80.6 (31.6)	
Pin-prick score, n	272	149	123	< 0.001*
Mean (SD)	73.6 (32.2)	68.5 (31.4)	79.9 (32.2)	
Type of paralysis	270	147	123	0.032†
Tetraplegic, n (%)	141 (52.2)	68 (46.3)	73 (59.3)	
Median UEMS in tetraplegic patients (IQR)	20.0 (10.0 to 36.0)	19.0 (8.0 to 36.0)	21.5 (10.0 to 38.5)	
Paraplegic, n (%)	129 (47.8)	79 (53.7)	50 (40.7)	
Traumatic lesion in spine, n (%)§	294	159	135	
Cervical spine	150 (51.0)	69 (43.4)	81 (60.0)	0.005†
Thoracic spine	109 (37.1)	73 (45.9)	36 (26.7)	0.001†
Lumbar spine	71 (24.1)	40 (25.2)	31 (23.0)	0.661†
Sacral spine	3 (1.0)	1 (0.6)	2 (1.5)	0.596‡
Time between injury and the first decompression, hrs, n¶	293	159	134	N/A
Mean (SD)	23.3 (36.3)	7.0 (2.6)	42.5 (46.9)	
Approach used for decompression, n (%)	282	151	131	0.139†
Anterior only	76 (27.0)	34 (22.5)	42 (32.1)	
Posterior only	190 (67.4)	110 (72.8)	80 (61.1)	
Circumferential only	9 (3.2)	3 (2.0)	6 (4.6)	
Combination of the above approaches	7 (2.5)	4 (2.6)	3 (2.3)	
Laminectomy, n (%)	282	151	131	0.039†
No	111 (39.4)	51 (33.8)	60 (45.8)	
Yes	171 (60.6)	100 (66.2)	71 (54.2)	

*Mann-Whitney U test.
†Chi-squared test.
‡Fisher's exact test.
§Patients could have traumatic lesions in multiple spine regions. Percentages do not add up to 100%.
¶If exact time of the first decompression was not reported but the quarter during surgery when decompression happened was available, the midpoint of the respective time interval was imputed. This variable is missing in one patient in whom the time of depression was unknown, and the patient was assigned to the late group based on the date of the decompression.
AIS, American Spinal Injury Association Impairment Scale; ASIA, American Spinal Injury Association; CCI, Charlson Comorbidity Index; IQR, interquartile range; LEMS, Lower Extremity Motor Score; N/A, not applicable; SD, standard deviation; UEMS, Upper Extremity Motor Score.

themselves highlight the low quality of available evidence, underpin the need for further research in this field.

To this end, the Prospective, Observational European Multicentre study on the efficacy of acute surgical decompression after traumatic Spinal Cord Injury (SCI-POEM) study was undertaken. The aim was to determine whether early (≤ 12 hours) surgical spinal decompression resulted in better neurological recovery 12 months after injury compared with late (> 12 hours and < 14 days) surgical spinal decompression in patients with acute tSCI.

Methods

Study design and patients. This multicentre prospective cohort study recruited patients from 17 centres in Europe. Patients were followed up for 12 months. The study was conducted between March 2013 and July 2019 (ClinicalTrials.gov identifier: NCT01674764).

Patients with tSCI presenting to the centres were assessed for eligibility against a predefined set of criteria. For inclusion to the study, patients had to fulfil all inclusion criteria and none of the exclusion criteria. Full inclusion and exclusion criteria are provided in Supplementary Table i. Major inclusion criteria were: age ≥ 18 years; diagnosis of blunt spinal column injury and SCI, including conus medullaris and/or cauda equina injuries; pre-surgery American Spinal Injury Association (ASIA) Impairment Scale (AIS) grade A, B, C, or D;[10] and < 14 days between injury and surgical spinal decompression. Major exclusion criteria included: traumatic brain injury with Glasgow Coma Scale ≤ 13; diagnosis of subclinical or clinical polyneuropathy; SCI caused by penetrating injury; non-traumatic or pathological fractures or cord compression; inability to cooperate with preoperative physical examination because of cognitive impairment; and previous spinal column or SCI.

The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines.[11,12] Ethics approval was obtained from each study centre before the beginning of patient enrolment. All patients provided written informed consent.

Study treatment groups. Depending on the timing of surgical spinal decompression, patients were divided into early (≤ 12 hours after injury) and late (> 12 hours and < 14 days after injury) decompression groups (early vs late group). The timing of surgical spinal decompression depended on the time elapsed between the injury and the patient's arrival at the study centre, the time for diagnostic investigations, and the judgment of the treating spinal surgeon. The specifics of the decompression surgery and stabilization (e.g. approach and the number of levels decompressed) as well as postoperative care were at the discretion of the treating spinal surgeon as per the standard of care of the study centre.

Neurological outcomes. Neurological examination was performed at baseline (after injury but before the decompression) according to the standards established by the ASIA or International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) examination (version 1.1).[10] Neurological deficits were assessed by the AIS grade, total motor score (MS), Upper Extremity Motor Score (UEMS), Lower Extremity Motor Score (LEMS), light touch score, and pin-prick score.[10] Scores were calculated using the ISNCSCI calculator to minimize rater bias.[13] Tetraplegic patients were those with the neurological level of injury at T1 or above, and paraplegic patients were those with the neurological level of injury at T2 or below. To assess the neurological recovery, ASIA examination was repeated at three-, six-, and 12-month follow-up.

Other study variables and safety outcomes. At baseline, patients' age, sex, Charlson Comorbidity Index (CCI),[14] Injury Severity Score (ISS),[15] and prior admission to another hospital after injury were collected. Patients completed the Spinal Cord Independency Measure (SCIM)[16] and World Health Organization Quality of Life (WHOQOL)-BREF at six and 12 months.[17] Scores of SCIM and WHOQOL-BREF were not the study outcomes of interest of this paper but were used for multiple imputation of missing data.

All adverse events (AEs) and serious AEs (SAEs) occurring during the 12-month follow-up were recorded and categorized into local (S)AEs (i.e. (S)AEs affecting the spine region), and general (S)AEs (i.e. (S)AEs affecting the rest of the body).

Study endpoints. The primary endpoint was the change in LEMS from baseline to 12 months. The secondary endpoints were the change in AIS grade from baseline to 12 months and the risks of developing at least one (S)AE/local (S)AE/general (S)AE during the follow-up.

Sample size estimation. Sample size estimation was based on the primary endpoint. We assumed a six-point difference in the change in LEMS between the two groups, which was considered the smallest clinically important difference. Based on available data from literature, an initial sample size calculation yielded a recruitment target of 300 patients. An internal pilot study comprising the first 100 patients who completed the 12-month follow-up, was used to update the estimates for the standard deviation (SD) of the change in LEMS, the allocation ratio, and the dropout rate, while the assumption of a six-point difference remained unchanged. Using a SD of 15.5, a 1:1 allocation ratio, an α level of 0.05, and a power of 0.8 for a two-sided independent-samples *t*-test, the updated sample size estimate was 212 patients. Assuming a rate of dropout or incomplete data of 30%, 303 patients were needed.

Statistical analysis. Statistical analyses were performed in all enrolled patients who received surgical spinal decompression with or without stabilization surgery. Baseline variables were compared between the two groups using independent-samples *t*-test, Mann-Whitney U test, chi-squared test, and Fisher's exact test, as appropriate. The between-group difference in the change in LEMS was calculated as the early group minus the late group. The change in LEMS from baseline to 12 months was compared between the two groups using both independent-samples *t*-test in a complete-case analysis (i.e. comprising patients with the outcome available at both baseline and 12 months) and analysis of covariance (ANCOVA) using multiply imputed data. The ANCOVA using multiply imputed data was the primary analysis of the study (confirmatory). Covariates in the ANCOVA comprised baseline LEMS, baseline AIS grade, and propensity score. The propensity score is the conditional probability of receiving a treatment of interest given a set of observed variables and has been used in non-randomized observational studies to address bias and confounding caused by imbalance in baseline

Table II. Baseline Lower Extremity Motor Score by baseline American Spinal Injury Association Impairment Scale grades and types of paralysis. Results are presented using patients with data available for individual variables; no imputation was performed to impute missing data.

Score	Early decompression (n = 159)	Late decompression (n = 135)	p-value*
LEMS by baseline AIS grades, n	149	124	
A, n (%)	72 (48.3)	38 (30.6)	0.850
Mean (SD)	1.9 (6.5)	3.1 (9.8)	
B, n (%)	34 (22.8)	12 (9.7)	0.904
Mean (SD)	6.2 (12.8)	7.5 (15.4)	
C, n (%)	19 (12.8)	22 (17.7)	0.338
Mean (SD)	10.4 (10.8)	12.6 (9.2)	
D, n (%)	24 (16.1)	52 (41.9)	0.291
Mean (SD)	40.5 (11.3)	43.3 (8.4)	
LEMS by types of paralysis†	147	123	< 0.001
Tetraplegic, n (%)‡	68 (46.3)	73 (59.3)†	
Mean (SD)	10.5 (17.5)‡	25.3 (21.7)	
Paraplegic, n (%)§	79 (53.7)	50 (40.7)	0.121
Mean (SD)	10.4 (16.2)	16.7 (18.7)§	

*Mann-Whitney U test.
†Numbers (%) of patients correspond to those presented in Table I. However, calculating mean (SD) LEMS for each subgroup does not comprise tetraplegic or paraplegic patients with missing data on baseline LEMS. Numbers of patients with missing data on baseline LEMS are indicated in the footnotes.
‡Based on 67 tetraplegic patients in the early group; one tetraplegic patient in the early group had missing baseline LEMS.
§Based on 47 paraplegic patients in the late group; three paraplegic patients in the late group had missing baseline LEMS.
AIS, ASIA Impairment Scale; ASIA, American Spinal Injury Association; LEMS, Lower Extremity Motor Score; SD, standard deviation; UEMS, Upper Extremity Motor Score.

variables between treatment groups.[18] Variables for generating the propensity score were age, sex, baseline AIS grade, baseline LEMS, baseline UEMS, CCI, ISS, type of injury (high- vs low-energy), location of the injury, and the number of levels affected. The balance diagnostics revealed a considerable residual imbalance in baseline AIS grade, i.e. the confounding effect of baseline AIS grade persisted after propensity-score adjustment, and could not be improved by any modification of the propensity score model. Therefore, we included baseline AIS grade as an additional covariate in all propensity score-adjusted analyses.

Missing data were expected and anticipated to be not completely at random; therefore, the unadjusted complete-case analysis, which only comprised patients with the outcome available at baseline and 12-month follow-up, would likely have led to biased conclusions. To accommodate for the systematic differences between observed and missing data, we used a fully conditional specification predictive mean matching method, which comprised a range of baseline and follow-up variables potentially predictive of the outcome, to impute missing LEMS.[19] The final imputation model comprised the following variables: all variables in the propensity score model, LEMS and UEMS at all follow-up visits, SCIM mobility subscale, and all WHOQOL-BREF domain scores at six months and 12 months, and occurrence of any SAEs or non-SAEs during the 12-month follow-up. When combining multiple imputation and propensity score analysis, the 'within approach' was applied,[20] in which (after imputing missing values multiple times and deriving the corresponding propensity score in each resulting dataset) the effect of the treatment group was estimated for each imputation and, finally, estimates were combined to produce an overall estimate.

Ancillary analyses of subgroup effect (interaction) were performed to investigate if the difference in the change in LEMS between the early and late group was dependent on the baseline AIS grade, the type of paralysis using a two-way analysis of variance without adjustment for covariates, and a two-way ANCOVA with adjustment for the covariates, i.e. 1) propensity score and baseline LEMS when analysing the subgroup effect of baseline AIS grade, and 2) propensity score, baseline LEMS, and baseline AIS grade when analyzing the subgroup effect of type of paralysis. The crude risk of developing at least one (S)AE during the 12 months was compared between the two groups using Fisher's exact test. The odds of experiencing at least one AE were compared between the treatment groups using a logistic regression model adjusting for propensity score and baseline AIS grade and the p-value was derived from the Wald test.

All statistical analyses were performed using SAS (version 9.4, SAS Institute, USA). All statistical tests were two-tailed. A p-value < 0.05 was deemed statistically significant. The final study protocol and the statistical analysis plan were signed on 19 October 2012, and 17 June 2020, respectively, rendering a previously published 'protocol' publication obsolete.[21]

Role of the funding source. AO Spine Europe and Southern Africa was the financial sponsor, supported the study design, and had the overall oversight of the study as per Good Clinical Practices. Study planning and execution, including data collection, study management, etc. were conducted by AO Innovation Translation Centre. The corresponding author (AJFH) had final responsibility for the decision to submit for publication.

Results

Disposition of patients. Of the 321 patients whose data were entered into the study database, 25 patients were excluded due to ineligibility, no written informed consent, no surgery done, or study site misconduct/non-compliance with data monitoring (Figure 1). Two patients were excluded as it could not

Table III. Descriptive change of American Spinal Injury Association Lower Extremity Motor Score (LEMS) between baseline and 12 months. Complete-case analysis comprised patients with both preoperative and 12-month follow-up LEMS assessed.

LEMS	Early decompression (n = 159)			Late decompression (n = 135)			Mean between-group difference (95% CI)	p-value
	n	Mean (95% CI)	Mean change from baseline (95% CI)	n	Mean (95% CI)	Mean change from baseline (95% CI)		
Unadjusted complete-case analysis								
Preoperative	109	11.2 (7.9 to 14.5)	N/A	86	24.1 (19.7 to 28.5)	N/A	N/A	N/A
12-mth follow-up	109	26.8 (22.6 to 31.0)	15.6 (12.1 to 19.0)	86	35.4 (31.3 to 39.6)	11.3 (8.3 to 14.3)	4.3 (-0.3 to 8.8)	0.065*
Adjusted analysis using multiple imputation								
12-mth follow-up	159	N/A	13.9 (11.3 to 16.4)	135	N/A	11.6 (9.1 to 14.2)	2.2 (-1.5 to 5.9)	0.245†

*Independent-samples *t*-test.
†Analysis of covariance using multiply imputed data comparing the change in LEMS between the two groups adjusting for baseline LEMS, baseline ASIA Impairment Scale grade, and propensity score. Presented values are least squares (LS) means and difference of the LS means using observed margins.
ASIA, American Spinal Injury Association; CI, confidence interval; LEMS, Lower Extremity Motor Score; N/A, not applicable.

be determined whether the surgical spinal decompression occurred ≤ 12 hours or > 12 hours after injury. The final analyses comprised 159 patients in the early group and 135 in the late group, who were recruited from 15 out of the 17 participating centres. The complete-case analysis of the change in LEMS, i.e. comprising patients with the outcome available at both baseline and 12 months, consisted of 109 patients in the early and 86 in the late group.

Baseline characteristics. Compared with those in the early group, patients in the late group were significantly older (p = 0.049, Mann-Whitney U test), and had lower mean ISS and a higher proportion of prior admission to another hospital after injury (Table I). Distributions of sex and CCI scores were similar. The distribution of baseline AIS grades differed significantly between the two groups, with a higher proportion of grade A in the early group (47.1% (n = 72) vs 29.2% (n = 38)) and a higher proportion of grade D in the late group (43.8% (n = 57) vs 17.6% (n = 27), Table I). The two groups differed significantly in baseline total MS, LEMS, light touch score, and pin-prick score, with greater scores in the late group. The distribution of the type of paralysis was also significantly different between the two groups. Baseline LEMS by AIS grade and type of paralysis are presented in Table II.

Within the early group, there were no significant differences in baseline characteristics between patients who completed (n = 119) and those who did not complete (n = 40) the follow-up (Supplementary Table ii). Within the late group, patients who completed (n = 93) and those who did not complete (n = 42) the follow-up differed significantly in baseline AIS grades, total MS, light touch score, and pin-prick score (Supplementary Table ii).

Changes in LEMS and analyses of subgroup effects. The mean change in LEMS from baseline to 12 months in the unadjusted complete-case analysis (i.e. comprising patients with the outcome available at both timepoints) was 15.6 (95% confidence interval (CI) 12.1 to 19.0) in the early group and 11.3 (95% CI 8.3 to 14.3) in the late group (Table III). The mean between-group difference in the change in LEMS was 4.3 (95% CI -0.3 to 8.8). Using the multiply imputed data, after adjusting for baseline LEMS, baseline AIS grade, and propensity score, the mean difference in the change in LEMS decreased to 2.2 (95% CI -1.5 to 5.9).

Both using a complete-case approach (i.e. comprising patients with all relevant variables available), results from the ancillary, covariate-adjusted analyses of subgroup effect (interaction) of the baseline AIS grade and type of paralysis (Table IV) were consistent with those without covariate adjustment (Supplementary Table iii). The propensity score and baseline LEMS-adjusted mean difference in the change in LEMS between the early and late group ranged from -1.3 (95% CI -12.1 to 9.5) for patients of baseline grade C to 4.6 (95% CI -3.1 to 12.3) for patients of baseline grade A (Table IV). The effect of treatment group was also greater in tetraplegic patients than in paraplegic patients (Table IV). However, there were no significant interactions between baseline AIS grade and treatment group and between baseline type of paralysis and treatment group, indicating that differences in the change in LEMS between the early and late group did not differ significantly between baseline AIS grades or between types of paralysis.

Change in AIS grade. Supplementary Table iv shows the change in AIS grade from baseline to 12 months with six-month observations carried forward. In this ancillary analysis, proportions of patients having improved (at least one grade), stabilized, and worsened (at least one grade) AIS grade were 53.7% (66/123), 39.8% (49/123), and 6.5% (8/123), respectively, in the early group and 50.5% (51/101), 46.5% (47/101), and 3.0% (3/101), respectively, in the late group. In patients of baseline AIS grades A to C, 39.2% (40/102) in the early group and 29.6% (16/54) in the late group had an improvement of ≥ 2 grades at 12 months.

Adverse events. A total of 221 AEs were recorded in 109/294 (37.1%) patients during the 12-month follow-up (Table V). The three most common AEs were urinary tract infection (61), respiratory AEs (34), and pressure ulcer (14). Six patients, three in each group, died due to a SAE (including one each due to renal failure, sepsis, cardiac event, and respiratory tract infection, and two due to other general SAE).

The crude risk of having at least one AE was significantly higher in the early than in the late group (43.4% (n = 69) vs 29.6% (n = 40) (Table V) unadjusted odds ratio (OR) comparing

Table IV. Analyses of subgroup effect (interaction) of baseline American Spinal Injury Association grade and type of paralysis on the difference in change in Lower Extremity Motor Score (between baseline and 12 months) between the early and late group (complete-case analyses, two-way analysis of covariance with adjustment of covariates). Complete-case analysis comprised patients with all relevant variables available.

Subgroup	Early decompression		Late decompression		Mean between-group difference (95% CI)	p-value
	n	Mean change from baseline (95% CI)	n	Mean change from baseline (95% CI)		
Baseline AIS grade*						0.788†
A	45	0.7 (-4.3 to 5.6)	16	-3.9 (-11.0 to 3.1)	4.6 (-3.1 to 12.3)	
B	28	21.6 (15.8 to 27.4)	9	17.4 (8.8 to 26.1)	4.2 (-6.0 to 14.3)	
C	10	23.8 (15.7 to 32.0)	13	25.1 (17.7 to 32.5)	-1.3 (-12.1 to 9.5)	
D	19	19.1 (11.7 to 26.5)	38	18.1 (11.7 to 24.6)	1.0 (-6.2 to 8.1)	
Type of paralysis‡						0.651§
Tetraplegic	45	16.7 (12.9 to 20.5)	45	14.0 (9.7 to 18.2)	2.7 (-2.9 to 8.4)	
Paraplegic	55	11.4 (7.5 to 15.3)	29	10.5 (5.7 to 15.2)	0.9 (-5.4 to 7.2)	

*Results from a two-way analysis of covariance with 1) change in LEMS as the dependent variable, 2) treatment group and baseline AIS grade as the two independent (categorical) variables), 3) an interaction term between baseline AIS grade and treatment group, and 4) baseline LEMS and propensity score as covariates. Presented values are least squares (LS) means and difference of LS means using observed margins.
†Interaction between treatment group and baseline AIS grade.
‡Results from a two-way analysis of covariance with 1) change in LEMS as the dependent variable, 2) treatment group and type of paralysis as the two independent (categorical) variables, 3) an interaction term between type of paralysis and treatment group, and 4) baseline LEMS, propensity score, and baseline AIS grade as covariates. Presented values are LS means and difference of LS means using observed margins.
§Interaction between treatment group and type of paralysis at baseline.
AIS, American Spinal Injury Association Impairment Scale; CI, confidence interval; LEMS, Lower Extremity Motor Score.

the early vs late group: 1.82 (95% CI 1.12 to 2.96)). After adjusting for propensity score and AIS at baseline, the odds of experiencing at least one AE did not differ significantly between the two groups (adjusted OR 1.38 (95% CI 0.75 to 2.56); p = 0.303, logistic regression model using Wald test). The crude risk of having at least one general AE was significantly higher in the early group (37.1% (n = 59) vs 25.2% (n = 34)), whereas the crude risk of having at least one local AE affecting the spine did not differ significantly between the two groups (14.5% (n = 23) vs 9.6% (n = 13)). The crude risks of having at least one SAE, local SAE, or general SAE were not significantly different between the two groups.

Discussion

The SCI-POEM study represents the largest prospective, multicentre study comparing early (≤ 12 hours) versus late (> 12 hours) surgical decompression for acute tSCI. Results from the unadjusted complete-case analysis (i.e. including patient with the outcome available at both baseline and 12 months) indicate a borderline significant difference, favouring the early group, in the mean change in LEMS from baseline to 12 months. However, after propensity score adjustment, early surgical spinal decompression was not associated with a statistically significant improvement at 12 months after injury. The higher incident rate of AEs seen in the early group was not considered to be associated with the timing of surgical intervention but rather attributed to the more severe neurological impairments seen in this group. This is supported by the results from our adjusted analysis in which after controlling for propensity score and baseline AIS grade, no significant differences between the two groups were found.

Our findings seem to be in concordance with those from the Surgical Trial in Acute Spinal Cord Injury Study (STASCIS).[22] Although the STASCIS study originally reported a positive association between early (≤ 24 hours) surgical decompression and improved neurological outcomes in patients with acute cervical tSCI, further dataset analyses showed different results.[23,24] First, a methodologically correct re-analysis of the 'two-grade AIS improvement' dataset did not find a statistically significant improvement at six months.[24] Second, by pooling the previously unpublished regression analysis data from the STASCIS study, a recent analysis showed statistically significant improvements neither in the total motor score (mean difference: 2.6 (95% CI -1.8 to 7.0); p = 0.24) nor in any other reported sensorimotor outcomes between early and late surgery group.[23] While differences between the STASCIS and SCI-POEM study exist, neither succeeded in confirming the century-old hypothesis, i.e. early surgical spinal decompression results in better neurological recovery than late surgical decompression in patients with acute tSCI.[3]

Due to the low incidence rate of tSCI, the logistical challenges faced by emergency medical services in the early hours after tSCI, and the profound lack of equipoise, an adequately powered, surgical randomized controlled trial is generally considered not feasible.[1,6,8,25] Our non-randomized study showed that in contrast to the early group, the majority of the late group were admitted to another hospital before admission to the study centres. In addition, patients in the early group had a higher rate of motor complete lesions (AIS grades A and B), as opposed to those in the late group having a higher rate of motor incomplete lesions (AIS grades C and D). Similar imbalances have been observed in other studies and may well be explained by the higher rate of indirect admissions to study centres among patients with less severe injuries.[26] This seems to override spinal surgeons' strong preference for providing surgical decompression to patients with an incomplete tSCI earlier than to those with a complete tSCI.[8] These difficulties, combined with the highly heterogenous presentation of acute tSCI, due to the possible range of affected anatomical and severity levels, pose various methodological challenges

Table V. Summary of adverse events occurring during the 12-month follow-up.

Event	Total (n = 294)		Early decompression (n = 159)		Late decompression (n = 135)		p-value‡
	n*	Risk, % (95% CI)†	n*	Risk, % (95% CI)†	n*	Risk, % (95% CI)†	
Any AE	109	37.1 (31.5 to 42.9)	69	43.4 (35.6 to 51.5)	40	29.6 (22.1 to 38.1)	0.016
Any serious AE	61	20.7 (16.3 to 25.8)	37	23.3 (16.9 to 30.6)	24	17.8 (11.7 to 25.3)	0.312
Any local AE (affecting the part of the body under investigation)	36	12.2 (8.7 to 16.5)	23	14.5 (9.4 to 20.9)	13	9.6 (5.2 to 15.9)	0.218
Any local serious AE	16	5.4 (3.1 to 8.7)	12	7.5 (4.0 to 12.8)	4	3.0 (0.8 to 7.4)	0.121
Postoperative bleeding	2	0.7 (0.1 to 2.4)	1	0.6 (0.0 to 3.5)	1	0.7 (0.0 to 4.1)	1.000
Dural tear	5	1.7 (0.6 to 3.9)	5	3.1 (1.0 to 7.2)	0	0.0 (0.0 to 2.7)	0.065
Cerebrospinal fluid leakage	3	1.0 (0.2 to 3.0)	1	0.6 (0.0 to 3.5)	2	1.5 (0.2 to 5.2)	0.596
Neurological deterioration	2	0.7 (0.1 to 2.4)	2	1.3 (0.2 to 4.5)	0	0.0 (0.0 to 2.7)	0.502
Hardware-related AEs§	6	2.0 (0.8 to 4.4)	3	1.9 (0.4 to 5.4)	3	2.2 (0.5 to 6.4)	1.000
Infections¶	3	1.0 (0.2 to 3.0)	2	1.3 (0.2 to 4.5)	1	0.7 (0.0 to 4.1)	1.000
Dysphagia	2	0.7 (0.1 to 2.4)	0	0.0 (0.0 to 2.3)	2	1.5 (0.2 to 5.2)	0.210
Pressure ulcers	12	4.1 (2.1 to 7.0)	5	3.1 (1.0 to 7.2)	7	5.2 (2.1 to 10.4)	0.395
Other local AE (affecting the part of the body under investigation)	8	2.7 (1.2 to 5.3)	6	3.8 (1.4 to 8.0)	2	1.5 (0.2 to 5.2)	0.296
Any general AE (affecting the rest of the body)	93	31.6 (26.4 to 37.3)	59	37.1 (29.6 to 45.1)	34	25.2 (18.1 to 33.4)	0.033
Any general serious AE	48	16.3 (12.3 to 21.1)	28	17.6 (12.0 to 24.4)	20	14.8 (9.3 to 21.9)	0.532
Cardiovascular AEs**	11	3.7 (1.9 to 6.6)	7	4.4 (1.8 to 8.9)	4	3.0 (0.8 to 7.4)	0.556
Respiratory AEs††	26	8.8 (5.9 to 12.7)	17	10.7 (6.4 to 16.6)	9	6.7 (3.1 to 12.3)	0.303
Urinary tract infection	44	15.0 (11.1 to 19.6)	28	17.6 (12.0 to 24.4)	16	11.9 (6.9 to 18.5)	0.191
Dehydration	1	0.3 (0.0 to 1.9)	0	0.0 (0.0 to 2.3)	1	0.7 (0.0 to 4.1)	0.459
Renal failure	1	0.3 (0.0 to 1.9)	0	0.0 (0.0 to 2.3)	1	0.7 (0.0 to 4.1)	0.459
Ileus	5	1.7 (0.6 to 3.9)	3	1.9 (0.4 to 5.4)	2	1.5 (0.2 to 5.2)	1.000
Sepsis	4	1.4 (0.4 to 3.4)	3	1.9 (0.4 to 5.4)	1	0.7 (0.0 to 4.1)	0.627
Other general AE (affecting the rest of the body)	38	12.9 (9.3 to 17.3)	19	11.9 (7.4 to 18.0)	19	14.1 (8.7 to 21.1)	0.605

The following AEs did not occur in any patients: intraoperative vascular injury, halo pin penetration, and dysphonia.
*Number of patients with at least one AE. If a patient experienced multiple AEs under any AE types, the patient was only counted once.
†Estimated crude risk of developing at least one AE (calculated by dividing the number of patients experiencing at least one AE by the total number of patients in the corresponding treatment group disregarding dropouts during the study). Confidence intervals calculated using the Clopper-Pearson method.
‡Fisher's exact test.
§Screw misplacement, loosening of instrumentation, and breakage of instrumentation (did not occur in any patient).
¶Surgical site infection and pin-track infection (did not occur in any patient).
**Cardiac events, pulmonary embolism, and deep vein thrombosis.
††Respiratory tract infection, acute respiratory distress syndrome, and atelectasis.
AE, adverse event; CI, confidence interval.

for the conduct and analysis of prospective, acute-stage surgical studies.[27]

We applied propensity score analysis, a well-validated statistical technique that allows for accurate assessment of treatment effects. Propensity score was used to adjust for patients' conditional probability of undergoing early or late surgical decompression according to a set of observed covariates. A recent study could not replicate the positive yet controversial findings on the effectiveness of methylprednisolone for the treatment of patients with tSCI from the Second National Spinal Cord Injury Study (NASCIS-II) study.[28,29] Instead, by applying propensity score-based adjustment, they found a significantly higher rate of total complications in the NASCIS-II methylprednisolone group, which led the authors not to recommend routine administration of methylprednisolone in acute tSCI.[29] Thus, in order to accurately estimate treatment effects in an adequately powered, yet inherently heterogeneous, cohort study of acute tSCI patients, sophisticated analytical techniques are warranted. Nonetheless, residual confounding, especially due to unobserved factors that could not be accounted for, still requires careful consideration.

The ISNCSCI AIS grade conversion rate is the most frequently reported neurological outcome measure used when investigating the timing of surgical decompression after acute tSCI.[30] However, the SCI-POEM study used LEMS improvement rather than AIS grade conversion as the primary outcome for the following reasons. First, AIS grade improvement does not equal motor score improvement. To illustrate, an improvement from AIS grade A to grade B does not require muscle strength improvement.[31] Second, AIS grade improvement does not always lead to functional recovery, which is particularly true for more severe injuries.[32] Third, the misclassification rates of AIS grades B and C are high, even for those with experience.[13] Fourth, performing accurate sacral sparing assessments can be challenging in the Emergency Room, which adds further to possible misclassification bias. Fifth, the AIS

grading system exhibits flooring and ceiling effects, particularly for two-grade improvements.[24] [24,33] However, in view of the six-point LEMS difference assumed for our sample size estimation, the late group exhibited a considerable LEMS ceiling effect due to the higher proportion of mildly injured patients seen in this group. To illustrate the potential impact of this hypothesized ceiling effect, we conducted a post-hoc sensitivity analysis on the primary endpoint including only patients with baseline LEMS < 45 points. In the unadjusted, complete-case analyses (i.e. the unadjusted analysis for patients with the outcome available at baseline and 12 months), the difference in LEMS change between patients undergoing early versus late surgical decompression dropped from 4.3 ((95% CI -0.3 to 8.8); p = 0.065, independent-samples *t*-test; Table III) to 1.6 (95% CI -3.5 to 6.7; p = 0.540, independent-samples *t*-test; Supplementary Table v). This sensitivity analysis indicates that observed ceiling effect partly contributed to an overestimate of the effect of early surgical decompression in the unadjusted analysis. In the adjusted analysis, we have attempted to limit the impact related to the described ceiling effect by including the baseline score as a covariate. In spite of inevitable ceiling effects seen in mildly injured tSCI subjects, the ISNCSCI UEMS and LEMS offer a more granular assessment of neurological impairment, have shown to be strongly correlated to functional outcomes, and have been recommended for use in interventional studies.[34] Furthermore, selecting LEMS improvement as a primary outcome also allowed the recruitment of both patients with tetraplegia and those with paraplegia, and eliminated the inherent variability of the UEMS seen in patients with tetraplegia.

While no consensus on threshold values pertaining to the minimally clinically important difference for LEMS is available,[35] the minimal detectable difference has shown to be in the range from one to seven motor score points, depending on the severity of the injury.[36] Thus, the adjusted mean difference of 2.2 LEMS points improvement favouring early surgical decompression seen in this study might actually not be clinically meaningful. Furthermore, variability in neurological examination timing within hours after acute tSCI may influence observations of long-term neurological recovery. Recent data from the Rick Hansen Spinal Cord Injury Registry demonstrated that patients of tSCI AIS grade A who were examined within eight hours after injury show three to four points higher total motor score improvement when compared to patients who underwent examination between eight and 48 hours after injury.[37] Therefore, one cannot exclude the possibility of the numerically higher motor score improvement seen in the early group being attributed to the timing of the initial neurological assessment rather than to the timing of the surgical spinal decompression. This confounder is, unfortunately, an inherent limitation for conducting a prospective observational interventional study where the timings of admission and examination are difficult to control.

Controversy remains about the role of early surgery in patients with traumatic central cord syndrome (TCCS). We did not include TCCS as a covariate in our analyses, since previous studies have shown that ISNCSCI/ASIA motor scores have a stronger predictive value than TCCS descriptors for neurological and functional outcomes in tSCI subjects.[38] Aarabi et al[39] confirmed that the extent of initial motor score deficit was a key prognostic factor for neurological recovery after TCCS, albeit in patients with evidence of pre-existing spinal stenosis. Age, however, was included as a covariate as this was a well-established, independent predictor of functional recovery following tSCI.[40] The presence of comorbidities may actually play a critical role in determining the optimal timing of spinal surgery and patient outcomes following TCCS.[41] Further, there is a common perception among surgeons that patients with traumatic cauda equina or conus medullaris syndromes exhibit a better recovery profile compared to patients with a traumatic injury to the spinal cord. We decided, however, to include all traumatic injuries to the spinal cord, conus medullaris, and cauda equina because there are no universally accepted definitions delineating these three 'regions' and, to the best of our knowledge, there is no unequivocal literature confirming the prognostic value of these three 'regions' on neurological recovery.

The strengths of the SCI-POEM study include the prospectively collected data from a large pan-European tSCI population with favourable follow-up rates, the availability of validated and detailed information on patients' pre-surgical neurological impairments assessed by trained and certified clinicians, the use of both a study protocol and a statistical analysis plan, the use of a validated ISNCSCI calculator to avoid misclassification, and the use of a propensity score adjustment.

Nonetheless, our study has several limitations. First, inherent to referral patterns seen in this heterogeneous group of acute tSCI patients, significant baseline differences were observed. These baseline imbalances in combination with earlier mentioned ceiling effects seen in mild tSCI may have yielded an overestimated borderline statistical significance favouring early surgical spinal decompression in unadjusted, complete-case analysis (i.e. the unadjusted analysis for patients with the outcome available at baseline and 12 months). A better understanding of any possible causal relationship between the severity of the injury and direct, or indirect, referral to a trauma centre is warranted. Second and unsurprisingly, a higher loss to follow-up rate was seen in more severely injured patients within the late group. This can be partly attributed to the inherent challenges of following severely injured patients who have been referred from, and reside in, rural communities. Third, while LEMS recovery is an established parameter of neurological recovery and strongly correlated to the ability to walk, other outcomes such as neurological level(s) of recovery, upper limb function as well as onset of recovery may also yield important information on potential therapeutic benefits. We collected data on these parameters as well as additional functional and quality of life outcomes, and plan to publish these in a second paper. Fourth, although our study had a single primary outcome which had been prespecified in the study protocol, we also conducted several sensitivity analyses, ancillary subgroup analyses, and secondary outcome analyses, e.g. AIS grade and AEs. Performing several tests on one dataset (multiple testing) is known to increase the type I error, if the significance level is not adjusted. Hence, results from our ancillary/subgroup analyses and secondary outcome analyses have to be regarded as exploratory in nature, and only the results from the propensity

score-adjusted analysis of the primary outcome, which we have used to draw our main conclusions, can be regarded as confirmatory. Fifth, postoperative imaging was not mandated for all patients to assess the extent of spinal cord decompression, which has been hypothesized to influence neurological recovery.[26] Sixth, there is no consensus on the threshold timing defining early versus late surgical decompression exists. The SCI-POEM study arbitrarily defined a threshold of 12 hours after injury, mainly driven by feasibility considerations and the pathophysiological rationale that mitigating secondary injury would be enhanced by performing surgery within this threshold. Hypothesis-generating analyses are planned in a second paper to assess the impact of the timing of surgical decompression using alternative thresholds (e.g. six or 24 hours after injury) as well as testing the impact of time as a continuous variable.[23] Finally, no standardized perioperative procedures or rehabilitation programmes were provided per the study protocol, these were tailored to the individual patient's needs and standard of care of the different centres. Nonetheless, despite the unknown extent of unobserved confounding, this multicentre observational study provided a unique framework to assess the potential benefit of acute treatment in the context of a contemporary, setting across the European region.

In conclusion, the SCI-POEM study has shown that early surgical decompression within 12 hours of injury did not result in statistically significant neurological improvements 12 months after injury. Studying interventions for acute tSCI is complicated by the inherent heterogeneity of this patient population, thus requiring sophisticated statistical techniques to reduce the impact of confounding variables. The application of such statistical analysis to control for this heterogeneity makes the results of SCI-POEM uniquely robust in the current SCI literature. Nonetheless, residual confounding, especially by unobserved factors that could not be accounted for, warrants careful consideration when interpreting our results and when drawing comparisons with data from other studies. These results do not impact the well-established need for acute medical treatment following tSCI, e.g. maintaining adequate spinal cord perfusion pressure. However, considering the medicolegal and operational challenges associated with implementing 'early surgical decompression' in current national trauma systems, we contend that these data warrant consideration when establishing clinical practice guidelines on the timing of surgical treatment for acute tSCI.[5]

Take home message

- Propensity score-adjusted analysis indicated that early (≤ 12 hours) surgical spinal decompression was not associated with a statistically significant neurological improvement at 12 months after injury.
- A combination of baseline imbalances, ceiling effects in mild traumatic spinal cord injury (tSCI), and a higher loss to follow-up rate of severe tSCI patients in the late treatment group may have yielded an overestimate of the effect of early surgical decompression in the unadjusted analysis.
- The SCI-POEM study is the first large comparative study in this area to account for heterogeneity in tSCI which results prompt critical reconsideration of international clinical practice guidelines on the timing of surgical treatment following tSCI.

Supplementary material

Tables of the inclusion and exclusion criteria, patient characteristics of study completers versus dropouts, subgroup analyses, and change in American Spinal Injury Association Impairment Scale grade between baseline and one year.

References

1. **Thietje R**, **Hirschfeld S**. Epidemiology of Spinal Cord Injury. In: Weidner N, Rupp R, Tansey KE, eds. *Neurological Aspects of Spinal Cord Injury*. Cham: Springer International Publishing, 2017: 3–17.
2. **Ahuja CS**, **Nori S**, **Tetreault L**, **et al**. Traumatic spinal cord injury-repair and regeneration. *Neurosurgery*. 2017;80(3S):S9–S22.
3. **Burrell HL**. I. Fracture of the spine: A summary of all the cases (244) which were treated at the Boston City Hospital from 1864 to 1905. *Ann Surg*. 1905;42(4):481–506.
4. **Batchelor PE**, **Wills TE**, **Skeers P**, **et al**. Meta-analysis of pre-clinical studies of early decompression in acute spinal cord injury: a battle of time and pressure. *PLoS One*. 2013;8(8):e72659.
5. **Fehlings MG**, **Tetreault LA**, **Wilson JR**, **et al**. A clinical practice guideline for the management of patients with acute spinal cord injury and central cord syndrome: Recommendations on the timing (≤ 24 hours versus > 24 hours) of decompressive surgery. *Global Spine J*. 2017;7(3 Suppl):195S–202S.
6. **van Middendorp JJ**, **Hosman AJF**, **Doi SAR**. The effects of the timing of spinal surgery after traumatic spinal cord injury: a systematic review and meta-analysis. *J Neurotrauma*. 2013;30(21):1781–1794.
7. **Ahuja CS**, **Badhiwala JH**, **Fehlings MG**. "Time is spine": the importance of early intervention for traumatic spinal cord injury. *Spinal Cord*. 2020;58(9):1037–1039.
8. **Fehlings MG**, **Rabin D**, **Sears W**, **Cadotte DW**, **Aarabi B**. Current practice in the timing of surgical intervention in spinal cord injury. *Spine (Phila Pa 1976)*. 2010;35(21 Suppl):S166–73.
9. **Rafter D**, **Vasdev R**, **Hurrelbrink D**, **et al**. Litigation risks despite guideline adherence for acute spinal cord injury: time is spine. *Neurosurg Focus*. 2020;49(5):E17.
10. **Kirshblum SC**, **Burns SP**, **Biering-Sorensen F**, **et al**. International standards for neurological classification of spinal cord injury (revised 2011). *J Spinal Cord Med*. 2011;34(6):535–546.
11. **World Medical Association**. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–2194.
12. **Committee for Human Medicinal Products**. Guideline for good clinical practice E6(R2) Step 5. European Medicines Agency. December 1, 2016. www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf (date last accessed 21 February 2023).
13. **Schuld C**, **Franz S**, **van Hedel HJA**, **et al**. International standards for neurological classification of spinal cord injury: classification skills of clinicians versus computational algorithms. *Spinal Cord*. 2015;53(4):324–331.
14. **Charlson ME**, **Pompei P**, **Ales KL**, **MacKenzie CR**. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–383.
15. **Baker SP**, **O'Neill B**, **Haddon W**, **Long WB**. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14(3):187–196.
16. **Catz A**, **Itzkovich M**, **Tesio L**, **et al**. A multicenter international study on the Spinal Cord Independence Measure, version III: Rasch psychometric validation. *Spinal Cord*. 2007;45(4):275–291.
17. **The WHOQOL Group**. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*. 1998;28(3):551–558.
18. **Rosenbaum PR**, **Rubin DB**. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc*. 1984;79(387):516–524.
19. **Barnard J**, **Rubin DB**. Miscellanea. Small-sample degrees of freedom with multiple imputation. *Biometrika*. 1999;86(4):948–955.
20. **Granger E**, **Sergeant JC**, **Lunt M**. Avoiding pitfalls when combining multiple imputation and propensity scores. *Stat Med*. 2019;38(26):5120–5132.
21. **van Middendorp JJ**, **Barbagallo G**, **Schuetz M**, **Hosman AJF**. Design and rationale of a Prospective, Observational European Multicenter study on the

efficacy of acute surgical decompression after traumatic Spinal Cord Injury: the SCI-POEM study. *Spinal Cord*. 2012;50(9):686–694.
22. **Fehlings MG**, **Vaccaro A**, **Wilson JR**, **et al**. Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS). *PLoS One*. 2012;7(2):e32037.
23. **Badhiwala JH**, **Wilson JR**, **Witiw CD**, **et al**. The influence of timing of surgical decompression for acute spinal cord injury: a pooled analysis of individual patient data. *Lancet Neurol*. 2021;20(2):117–126.
24. **van Middendorp JJ**. Letter to the editor regarding: "Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS)." *Spine J*. 2012;12(6):540.
25. **Ng WP**, **Fehlings MG**, **Cuddy B**, **et al**. Surgical treatment for acute spinal cord injury study pilot study #2: evaluation of protocol for decompressive surgery within 8 hours of injury. *Neurosurg Focus*. 1999;6(1):e3.
26. **Aarabi B**, **Akhtar-Danesh N**, **Chryssikos T**, **et al**. Efficacy of ultra-early (< 12 h), early (12-24 h), and late (> 24-138.5 h) surgery with magnetic resonance imaging-confirmed decompression in American Spinal Injury Association Impairment Scale grades A, B, and C cervical spinal cord injury. *J Neurotrauma*. 2020;37(3):448–457.
27. **Geisler FH**, **Coleman WP**, **Grieco G**, **Poonian D**, **Sygen Study G**. Measurements and recovery patterns in a multicenter study of acute spinal cord injury. *Spine (Phila Pa 1976)*. 2001;26(24 Suppl):S68–86.
28. **Bracken MB**, **Shepard MJ**, **Collins WF**, **et al**. A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury. *N Engl J Med*. 1990;322(20):1405–1411.
29. **Evaniew N**, **Noonan VK**, **Fallah N**, **et al**. Methylprednisolone for the treatment of patients with acute spinal cord injuries: A propensity score-matched cohort study from a Canadian multi-center spinal cord injury registry. *J Neurotrauma*. 2015;32(21):1674–1683.
30. **Wilson JR**, **Witiw CD**, **Badhiwala J**, **Kwon BK**, **Fehlings MG**, **Harrop JS**. Early surgery for traumatic spinal cord injury: Where are we now? *Global Spine J*. 2020;10(1 Suppl):84S–91S.
31. **van Middendorp JJ**, **Hosman AJF**, **Pouw MH**, **Van de Meent H**, **EM-SCI Study Group**. Is determination between complete and incomplete traumatic spinal cord injury clinically relevant? Validation of the ASIA sacral sparing criteria in a prospective cohort of 432 patients. *Spinal Cord*. 2009;47(11):809–816.
32. **van Middendorp JJ**, **Hosman AJF**, **Pouw MH**, **Van de Meent H**, **EM-SCI Study Group**. ASIA impairment scale conversion in traumatic SCI: is it related with the ability to walk? A descriptive comparison with functional ambulation outcome measures in 273 patients. *Spinal Cord*. 2009;47(7):555–560.
33. **Fawcett JW**, **Curt A**, **Steeves JD**, **et al**. Guidelines for the conduct of clinical trials for spinal cord injury as developed by the ICCP panel: spontaneous recovery after spinal cord injury and statistical power needed for therapeutic clinical trials. *Spinal Cord*. 2007;45(3):190–205.
34. **Marino RJ**, **Graves DE**. Metric properties of the ASIA motor score: subscales improve correlation with functional activities. *Arch Phys Med Rehabil*. 2004;85(11):1804–1810.
35. **Wu X**, **Liu J**, **Tanadini LG**, **et al**. Challenges for defining minimal clinically important difference (MCID) after spinal cord injury. *Spinal Cord*. 2015;53(2):84–91.
36. **Marino RJ**, **Jones L**, **Kirshblum S**, **Tal J**, **Dasgupta A**. Reliability and repeatability of the motor and sensory examination of the international standards for neurological classification of spinal cord injury. *J Spinal Cord Med*. 2008;31(2):166–170.
37. **Evaniew N**, **Sharifi B**, **Waheed Z**, **et al**. The influence of neurological examination timing within hours after acute traumatic spinal cord injuries: an observational study. *Spinal Cord*. 2020;58(2):247–254.
38. **Pouw MH**, **van Middendorp JJ**, **van Kampen A**, **Curt A**, **van de Meent H**, **Hosman AJF**. Diagnostic criteria of traumatic central cord syndrome. Part 3: descriptive analyses of neurological and functional outcomes in a prospective cohort of traumatic motor incomplete tetraplegics. *Spinal Cord*. 2011;49(5):614–622.
39. **Aarabi B**, **Akhtar-Danesh N**, **Simard JM**, **et al**. Efficacy of early (≤ 24 hours), late (25-72 hours), and delayed (> 72 hours) surgery with magnetic resonance imaging-confirmed decompression in American Spinal Injury Association Impairment Scale grades C and D acute traumatic central cord syndrome caused by spinal stenosis. *J Neurotrauma*. 2021;38(15):2073–2083.
40. **van Middendorp JJ**, **Hosman AJF**, **Donders ART**, **et al**. A clinical prediction rule for ambulation outcomes after traumatic spinal cord injury: A longitudinal cohort study. *Lancet*. 2011;377(9770):1004–1010.
41. **Samuel AM**, **Grant RA**, **Bohl DD**, **et al**. Delayed surgery after acute traumatic central cord syndrome is associated with reduced mortality. *Spine (Phila Pa 1976)*. 2015;40(5):349–356.

Author information:
A. J. F. Hosman, MD, PhD, Orthopedic Surgeon, Principal Investigator
J. J. van Middendorp, PhD, Medical Doctor, Clinical Researcher, Co-Principal Investigator
Department of Orthopaedic Surgery, Radboud University Medical Center, Nijmegen, Netherlands.

G. Barbagallo, MD, Neurosurgeon, U.O. di Clinica Neurochirurgia Azienda Ospedaliero, Universitaria Policlinico, Catania, Italy.

Author contributions:
A. J. F. Hosman: Conceptualization, Methodology, Writing –original draft, Writing – review & editing, Investigation, Supervision.
J. J. van Middendorp: Conceptualization, Methodology, Writing –original draft, Writing – review & editing, Investigation, Supervision.
G. Barbagallo: Funding acquisition, Writing – review & editing, Investigation, Supervision.

Funding statement:
The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: this study was financially supported by AO Spine Europe and Southern Africa. AO Spine is a clinical division of the AO Foundation, which is an independent medically guided not-for-profit organization.

ICMJE COI statement:
The authors have no conflicts of interest to declare.

Data sharing:
Deidentified participant data with corresponding data dictionary of the data underlying the current manuscript will be made available on reasonable request to the corresponding author after the publication of all secondary endpoints. Additional documents such as the study protocol and statistical analysis plan will also be available on request. The data will be shared with external researchers for scientific non-commercial purposes with investigator support after approval of the proposal by the AO Spine SCI-POEM steering board, including a signed data user agreement.

Acknowledgements:
Study management was performed by AO Innovation Translation Center (AO ITC), Clinical Evidence and Clinical Operations. The authors would like to thank all AO ITC staff for their assistance with the study, specifically, Brigitte S. Gallo-Kopf, Joffrey Baczkowski, Janik Hilse, Tracy Yaner Zhu, Anahi Hurtado, Bettina Krummenacher, Benjamin Weiss, Marina Meuwly, Nicole Varga, Julie Gujan, and Kathrin Espinoza-Rebmann. The authors, AO Spine, and AO ITC (former AOCID) would like to express their gratitude to all study site personnel for their continued support and efforts to the study. The authors thank Dr Christian Schuld, Dr Rüdiger Rupp, and Prof. Armin Curt for their support in utilizing the EMSCI's ISNCSCI calculator and Dr Rogier Donders for his advice.
The SCI-POEM Study Group:
Allard J. F. Hosman (Radboud University Medical Center, Nijmegen, The Netherlands), Giuseppe Barbagallo (Policlinico "G.Rodolico - San Marco" University Hospital, Catania, Italy), Eugen Cezar Popescu ("Prof. N. Oblu" Emergency Hospital, Iasi, Romania), Henk van de Meent (Radboud University Medical Centre, Nijmegen, The Netherlands), F. Cumhur Öner (Universitair Medisch Centrum Utrecht, Utrecht, The Netherlands), Federico De Iure (Ospedale Maggiore – AUSL Bologna, Italy), Jacopo Bonavita (Montecatone Rehab Institute, Spinal Unit, Imola, Italy), Michael Kreinest (BG Klinik Ludwigshafen, Ludwigshafen, Germany), Richard A. Lindtner (Medical University of Innsbruck, Innsbruck, Austria), Nasir A Quraishi (Nottingham University Hospital NHS Trust, Nottingham, UK), Pradeep Thumbikat (Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK), Vide Bilić (University Hospital Centre Sestre Milosrdnice, Zagreb, Croatia), Jeremy J Reynolds (Oxford University Hospitals NHS Trust, Oxford, UK), Maurizio Belci (National Spinal Injury Centre Stoke Mandeville Hospital, UK), Alp Özgün Börcek (Gazi Üniversitesi Tıp Fakültesi Beyin ve Sinir Cerrahisi Anabilim Dalı, Ankara, Turkey), Seamus Morris (Mater Misericordiae University Hospital / UCD School of Medicine, Dublin, Ireland), Christoph Hoffmann (Center for Spinal Surgery and Neurotraumatology, Frankfurt, Germany), Francesco Signorelli (Hôpital Neurologique et Neurochirurgical Pierre Wertheimer, Lyon, France), Konstantin Uzunov (University Multiprofile Hospital for Active Treatment and Emergency Medicien "N.I.Pirogov", Sofia, Bulgaria), Joost

J. van Middendorp (Radboud University Medical Center, Nijmegen, The Netherlands).

Open access funding
The open access fee for this study was funded by AO Spine Europe and Southern Africa.

Trial registration number:
ClinicalTrials.gov identifier: NCT01674764.

This article was primary edited by S. P. F. Hughes.